ACCURACY IMPROVEMENT OF HANDWRITTEN CHARACTER RECOGNITION BY GLVQ

Fukumoto, T.; Wakabayashi, T.; Kimura, F.; Miyake, Y.

ACCURACY IMPROVEMENT OF HANDWRITTEN CHARACTER RECOGNITION BY GLVQ

TSUYOSHI FUKUMOTO, TETSUSHI WAKABAYASHI, FUMITAKA KIMURA, and YASUJI MIYAKE

Faculty of Engineering, Mie University, 1515 Kamihama, Tsu 514-8507, JAPAN
E-mail: kimura@hi.info.mie-u.ac.jp

This paper deals with accuracy improvement of handwritten character recognition by the GLVQ (generalized learning vector quantization). In literature [3], the way of combining the FDA (Fisher discriminant analysis) and the GLVQ was investigated and evaluated to be effective for handwritten Chinese character recognition employing the minimum Euclidian distance classifier. In this paper, the projection distance and the modified projection distance are employed besides the Euclidian distance, and handwritten numerals as well as Chinese characters are used for the evaluation test. The result of experiment shows that the learning of reference vectors by GLVQ improves the recognition accuracy of not only the Euclidian distance classifier but also the projection distance classifier and the modified projection distance classifier. The highest accuracy (98.41%) for the Chinese character recognition was obtained when the FDA, GLVQ and the modified projection distance were employed. The highest accuracy (99.36%) for the numeral recognition was obtained when the GLVQ and the modified projection distance were employed.

1 Introduction

This paper deals with accuracy improvement of handwritten character recognition by the GLVQ (generalized learning vector quantization). The GLVQ is a generalization of the LVQ [8] and is formalized as a minimization problem of an evaluation function, which guarantees the convergence of the reference vectors [1]. A linear transformation of feature vector by the FDA (Fisher discriminant analysis) [6] efficiently improves the recognition accuracy of Chinese character set, which has a large number of character classes [2]. In literature [3], the way of combining the FDA and the GLVQ was investigated and evaluated to be effective for handwritten Chinese character recognition employing the minimum Euclidian distance classifier. In this paper, the projection distance and the modified projection distance are employed besides the Euclidian distance, and handwritten numerals as well as Chinese characters are used for the evaluation test.

2 GLVQ Algorithm

Proximity of an input feature vector X to its own class is defined by

$$\mu(X) = \frac{d_m(X) - d_j(X)}{d_m(X) + d_j(X)} \tag{1}$$

, where d_m is some distance between X and the reference vector R^m of the class to which X belongs, and d_j is the distance between X and the nearest reference vector R^j of the class to which X does not belongs. The proximity μ takes the value in [-1, 1] for any X, and leads to correct (incorrect) classification when it is negative (positive).

The purpose of the learning is to reduce the incorrect classification by reducing the value of μ for as many X as possible. The GLVQ is formalized as a minimization problem of an evaluation function Q defined by

$$Q = \sum_{i=0}^{N} f(\mu_i) \tag{2}$$

, where N is the total number of the input vectors, and $f(\mu)$ is a monotonically increasing function of μ. The evaluation function S can be minimized by iterative modification of the reference vectors

$$R^m_{k,t+1} = R^m_{k,t} + \alpha \cdot \frac{\partial f}{\partial \mu} \frac{d_j}{(d_m + d_j)^2} (X_k - R^m_{k,t}) \tag{3}$$

$$R^j_{k,t+1} = R^j_{k,t} - \alpha \cdot \frac{\partial f}{\partial \mu} \frac{d_m}{(d_m + d_j)^2} (X_k - R^j_{k,t}) \tag{4}$$

, which are derived by the steepest descent method. In the experiment, the first derivative of the function f is defined by

$$\frac{\partial f}{\partial \mu} = F(\mu, t)(1 - F(\mu, t)) \tag{5}$$

, where k is the dimensionality, t is the number of iteration, and $F(\mu, t)$ is the sigmoid function defined by

$$F(\mu, t) = \frac{1}{1 + e^{-\mu(X)t}} \tag{6}$$

As the initial reference vectors, the mean vector of each class are used.

3 Distance Function for Classification

3.1 Euclidian Distance

The Euclidian distance between the input pattern and the mean vector is defined by

$$g_l^2(X) = \|X - M_l\|^2 \tag{7}$$

, where X is the input feature vector, M_l is the mean vector of class l. The input vector is classified to such class l^* that minimizes the Euclidian distance.

Hereafter the subscript l denoting the class is omitted for simplicity's sake.

3.2 Projection Distance

The projection distance [4] is defined by

$$g^2(X) = \|X - M\|^2 - \sum_{i=1}^{k} \{\Phi_i(X - M)\}^2 \tag{8}$$

and gives the distance from the input pattern X to the minimum square error hyperplane which approximates the distribution of the samples, where Φ_i denotes the i_{th} eigenvector of the covariance matrix, and k is the dimensionality of the hyperplane. Fig.1 shows an example of decision boundary in two dimensional feature space. This figure shows that the first principal axis ap-

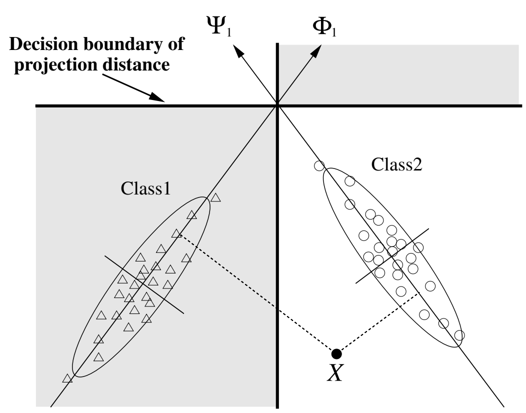

Figure 1: Decision boundary of projection distance (2-dimensional)

proximates the distribution with minimum square error, and the distance from an input pattern X to the axis determines the class. In this figure the decision boundary consists of a pair of lines (asymptotic lines of a hyperbola) which is degenerated special case of a quadratic curve. In general the decision boundary consists of quadratic hypersurfaces.

3.3 Modified Projection Distance

The increase of the classification error in and near the intersecting point of the hyperplanes (shared subspace) has been pointed out as a drawback of the projection distance method (subspace method) [5]. If Ex. (8) is modified as

$$g^2(X) \quad = \quad \|X - M\|^2 - \sum_{i=1}^{k} \frac{(1-\alpha)\lambda_i}{(1-\alpha)\lambda_i + \alpha\sigma^2} \{\Phi_i(X - M)\}^2 \qquad (9)$$

the classification error due to the subspace intersection can be suppressed, where α is a parameter [-1, 1], σ^2 is the average of the all eigen values of all classes. The modified projection distance with $\alpha = 0$ is equivalent to the projection distance, and the one with $\alpha = 1$ is to the Euclidian distance.

The hyperbola in Fig.2 is an example of the decision boundary of the modified projection distance. While the projection distance method yields incorrect classification from class 2 to 1 around the intersection of the hyperplanes (lines in this example) because the distance between the input pattern and the mean vector along the hyperplane is completely neglected, the modified projection

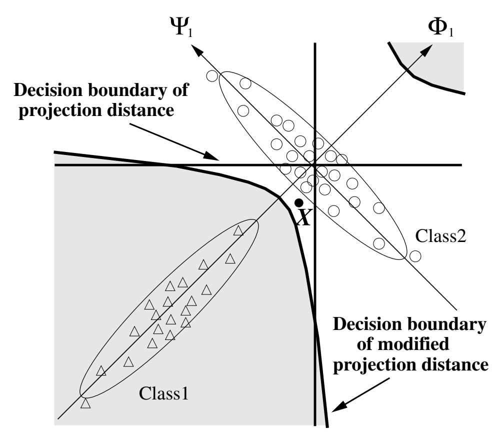

Figure 2: Decision boundary of modified projection distance (2-dimensional)

distance takes the distance along the hyperplane into account, to some extent, and does not yields those misclassifications.

When the dimensionality of feature vector is large, and the dimensionality of the hyperplane is relatively small, the probability of intersection and hence the misclassification due to the subspace intersection decreases. In the practical application of the projection distance and the subspace method, such conditions are usually satisfied.

4 Fisher Discriminant Analysis

A linear transformation of feature vector by the FDA (Fisher discriminant analysis) [6] efficiently improves the recognition accuracy of Chinese characters, which have a large number of character classes [2].

The discriminant analysis selects an n' dimensional feature vector from an n dimensional feature vector as follows ($n' < n$)[6]. The eigenvalues and eigenvectors satisfying

$$S_b \Phi = S_w \Phi \Lambda \tag{10}$$

is calculated, where matrices Λ and Φ are eigenvalue matrix having eigenvalues λ_i ($\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$) as its diagonal elements, and eigenvectors Φ_i as its column vectors, respectively. The between-class covariance matrix S_b and within-class covariance matrix S_w are respectively defined by

$$S_w = \sum_{l=1}^{L} P(w_l) E\{(X - M_l)^T | w_l\} \tag{11}$$

$$S_b = \sum_{l=1}^{L} (M_l - M_0)(M_l - M_0)^T \tag{12}$$

, where L is the number of classes, $P(w_l)$ and M_l are the *a priori* probability and the mean vector of the class l. M_0 is the total mean vector and is defined by

$$M_0 = \sum_{l=1}^{L} P(w_l) M_l \tag{13}$$

When the rank of S_b is r, maximum of r nonzero eigen values are obtained. The rank r is less than or equal to $min(L-1, n)$.

By the linear combination for the feature vector X

$$y_i = \Phi_i^T X \qquad (i = 1, 2, ..., n') \tag{14}$$

n' Fisher discriminants with n' largest variance ratio (eigenvalue λ_i) are selected as new features. When $n' = n$, the new feature vector is given by

$$Y = \Phi^T X \tag{15}$$

5 Recognition Experiment

5.1 Handwritten Chinese Character Recognition

In this experiment, 3036 classes of handwritten Chinese characters in ETL9B collected by the Electrotechnical Laboratory were used. A feature vector of size 196 was extracted from each character image by the weighted direction code histogram method [2]. The sample size is 200 per class and was divided into learning sample and test sample of size 100 per class respectively. The performance evaluation of the three classifiers with the reference vectors obtained by the GLVQ as well as the original mean vectors was performed for comparison. The result is shown in Table.1. This table implies that

1. the feature transformation by the FDA improves the recognition accuracy,
2. the modification of the reference vectors by the GLVQ improves the recognition accuracy,
3. the direct cascading of the FDA and the GLVQ further improves the recognition accuracy,
4. the accuracies of the classifiers employing the Euclidian distance, the projection distance, and the modified projection distance are increasing in this order.

The number of iteration of the GLVQ is about 40 for the Euclidian distance classifier, and is about 20 for the other two classifies. When the FDA is applied, the size (dimensionality) of the transformed feature vector was 144.

While the Euclidian distance classifier is suitable for classification of spherically distributed patterns, the performance is deteriorated when the distribution is elliptical or asymmetric. While the performance of the projection distance classifier and the modified projection distance classifier is not deteriorated by the elliptical distribution, it is deteriorated by the asymmetric distribution. The accuracy improvement by the GLVQ is attributed to the elimination of these defects of classifiers by the optimization of the reference vectors. In the Chinese character recognition by the projection distance, the misclassification due to the subspace intersection is also reduced by the GLVQ.

Table 1: Result of handwritten Chinese character recognition

Distance	Transformation and Learning	Recognition accuracy(%)
Euclidian distance	-	94.61
	FLD	96.73
	GLVQ	97.22
	FLD + GLVQ	97.51
Projection distance	-	97.26
	FLD	98.05
	GLVQ	97.83
	FLD + GLVQ	98.30
Modified projection distance	-	98.02
	FLD	98.32
	GLVQ	98.32
	FLD + GLVQ	98.41

5.2 Handwritten Numeral Recognition

In this experiment, handwritten ZIP code numerals in IPTP CDROM1 collected by the Institute for Posts and Telecommunications Policy were used. A feature vector of size 400 was extracted from each numeral image by the gradient analysis method [2], [7]. The total sample was divided into learning sample and test sample of size 29,883 and 14,979 respectively. The performance evaluation of the three classifiers with the reference vectors obtained by the GLVQ as well as the original mean vectors was performed for comparison. In the numeral recognition (with ten classes), the FDA was not applied because of the rank deficiency of the between-class covariance matrix. Instead, the dimension reduction by the principal component analysis was applied to study the relationship between the recognition accuracy and the dimensionality.

The result is shown in Table.2 and Fig.3. These results imply that

1. the modification of the reference vectors by the GLVQ improves the recognition accuracy,
2. the accuracies of the classifiers employing the projection distance, and the modified projection distance are higher than the Euclidian distance classifier,
3. the feature size is reduced to 100 by the principal component analysis without sacrificing the recognition accuracy.

Fig.4 shows the relationship between the recognition accuracy and the number of the iteration of the GLVQ.

In the numeral recognition, the performance of the projection distance classifier and the modified projection distance classifier is almost the same. The performance of the projection distance is not deteriorated by the subspace intersection because the number of classes are small (ten).

Table 2: Result of handwritten numeral recognition

Dimensionality	Distance	test sample		learning sample	
		Original	GLVQ	Original	GLVQ
64	Euclidian distance	94.18	98.74	93.20	99.39
	Projection distance	99.12	99.14	99.06	99.15
	Modified projection distance	99.12	99.22	99.07	99.63
100	Euclidian distance	94.19	98.74	93.23	99.32
	Projection distance	99.17	99.22	99.20	99.40
	Modified projection distance	99.19	99.35	99.27	99.70
196	Euclidian distance	94.22	98.79	93.24	99.37
	Projection distance	99.22	99.32	99.32	99.48
	Modified projection distance	99.22	99.36	99.39	99.68
400	Euclidian distance	94.21	98.84	93.24	99.49
	Projection distance	99.23	99.34	99.34	99.58
	Modified projection distance	99.23	99.34	99.34	99.82

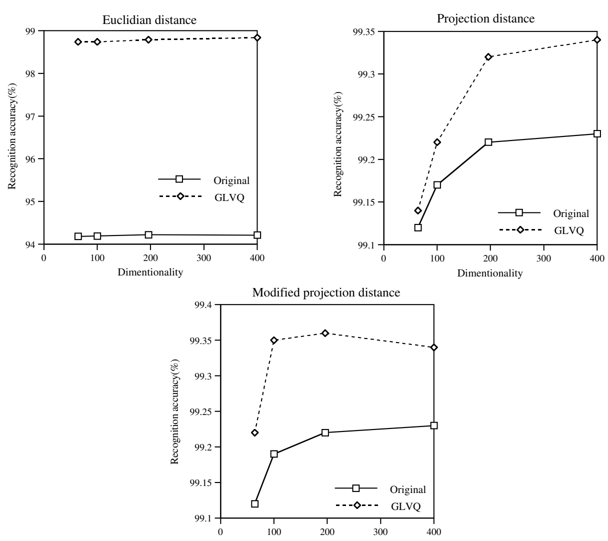

Figure 3: Recognition accuracy (for test data) v.s. dimensionality

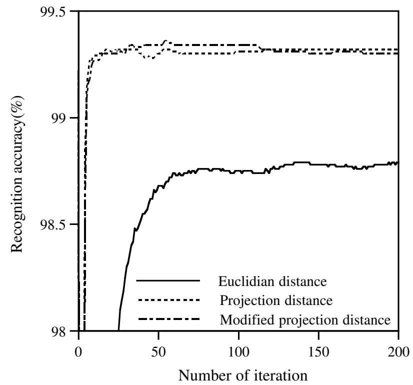

Figure 4: Recognition accuracy (for test data) v.s. number of iteration of GLVQ

6 Conclusion

The result of experiment showed that the learning of reference vectors by GLVQ improves the recognition accuracy of not only the Euclidian distance classifier but also the projection distance classifier and the modified projection distance classifier. The highest accuracy (98.41%) for the Chinese character recognition was obtained when the FDA, GLVQ and the modified projection distance were employed. The highest accuracy (99.36%) for the numeral recognition was obtained when the GLVQ and the modified projection distance were employed.

Studies on the accuracy improvement by learning multiple reference vectors per class, and learning of the feature transformation matrix, are remaining as interesting future research topics.

References

1. A.Sato and K.Yamada, "Generalized learning vector quantization", Advances in Neural Information Processing 8, Proc. of the 1995 Conference, pp.423-429, MIT Press, Cambridge, MA, USA, 1996.
2. F.Kimura, T.Wakabayashi, S.Tsuruoka and Y.Miyake, "Improvement of Handwritten Japanese Character Recognition Using Weighted Direction Code Histogram", Pattern Recognition, Vol.30, No.8, pp.1329-1337, 1997.
3. K-H.Shyu, M-K.Tsay and P-C.Chang, "Feature Transformation with Generalized Learning Vector Quantization for Hand-Written Chinese

Character Recognition", IEICE Transactions on Information and Systems, Vol.E82-D, No.3 , pp.687-692, 1999.

4. M.Ikeda, H.Tanaka and T.Moto-oka, "Projection distance method for recognition of hand-written characters (in Japanese), Trans. IPS Japan, Vol.24, No.1, pp.106-112, 1983.
5. Oja, E., "Subspace Method of Pattern Recognition", Reserch Studies Press, England, 1983.
6. K.Fukunaga, "Introduction to Statistical Pattern Recognition, Second Edition", Academic Press, 1990.
7. T.Wakabayashi, S.Tsuruoka, F.Kimura and Y.Miyake, "Increasing the feature size in handwritten numeral recognition to improve accuracy", Systems and Computers in Japan (English Edition), Scripta Technica, Inc. Vol.26, No.8, pp.35-44, 1995.
8. T.Kohonen, "Self-Organization and Associative Memory, 3rd ed.", Springer-Verlag, 1989.